

APPENDIX A: Document Type Definition of Multimedia Integration Description Scheme

**MM\_integration.ds.dtd**

5

```
<!-- Multimedia Integration Description Scheme -->
```

```
<!ELEMENT MM_Stream ( MM_Object_Set, Object_Hierarchy*,  
Entity_Relation_Graph* )>
```

10

```
<!ATTLIST MM_Stream  
    id ID #IMPLIED>
```

```
<!-- This is how external DTDs are included in the current DTD -->
```

15

```
<!-- External Video DS DTD -->
```

```
<!ENTITY % Video_DS SYSTEM "video.ds.dtd">  
%Video_DS;
```

```
<!--External Audio DS DTD -->
```

20

```
<!ENTITY % Audio_DS SYSTEM "audio.ds.dtd">  
%Audio_DS;
```

```
<!--External Text DS DTD -->
```

25

```
<!ENTITY % Text_DS SYSTEM "text.ds.dtd">  
%Text_DS;
```

```
<!--External Synthetic DS DTD -->
```

```
<!ENTITY % Synthetic_DS SYSTEM "synthetic.ds.dtd">  
%Synthetic_DS;
```

30

```
<!--External Image DS DTD -->
```

```
<!ENTITY % Image_DS SYSTEM "image.ds.dtd">  
%Image_DS;
```

35

```
<!ELEMENT MM_Object_Set ( MM_Object+ )>
```

```
<!ELEMENT MM_Object ( Media_Object_Set, Object_Hierarchy*,  
Entity_Relation_Graph*,
```

40

```
        MM_Obj_Media_Features?, MM_Obj_Semantic_Features?,  
        MM_Obj_Temporal_Features? )>
```

```
<!ATTLIST MM_Object
```

```
    Object_Type
```

```
(LOCAL|GLOBAL)
```

```
#REQUIRED
```

	<pre> id           ID          #IMPLIED Object_Ref   IDREF      #IMPLIED Object_Node_Ref IDREFS    #IMPLIED Entity_Node_Ref IDREFS    #IMPLIED&gt;  5 &lt;!ELEMENT Media_Object_Set ( Audio_Object   Image_Object   Video_Object   Text_Object                                 Synthetic_Object )+ &gt;  10 &lt;!-- The object hierarchy and the entity relation graph are defined in the Image DS (Proposal #480). We include them in this DTD for convenience --&gt; &lt;!-- Object hierarchy element --&gt; &lt;!-- The attribute type is the hierarchy binding type --&gt; &lt;!ELEMENT Object_Hierarchy ( Object_Node )&gt; 15 &lt;!ATTLIST Object_Hierarchy       id           ID          IMPLIED       type         CDATA      #IMPLIED&gt;  &lt;!ELEMENT Object_Node ( Object_Node* )&gt; 20 &lt;!ATTLIST object_node       id           ID          #IMPLIED       Object_Ref   IDREF      #REQUIRED&gt;  &lt;!-- Entity relation graph element--&gt; 25 &lt;!-- Possible types of entity relations and entity relation graphs:       - Spatial: topological, directional       - Temporal: topological, directional       - Semantic --&gt; &lt;!ELEMENT Entity_Relation_Graph ( Entity_Relation+ )&gt; 30 &lt;!ATTLIST Entity_Relation_Graph       id           ID          #IMPLIED       type         CDATA      #IMPLIED&gt;  &lt;!ELEMENT Entity_Relation ( Relation, (Entity_Node   Entity_Node_Set   Entity_Relation)* )&gt; 35 &lt;!ATTLIST Entity_Relation       type         CDATA      #IMPLIED&gt;  &lt;!ELEMENT Entity_Node (#PCDATA)&gt; 40 &lt;!ATTLIST Entity_Node       id           ID          #IMPLIED       Object_Ref   IDREF      #REQUIRED&gt;  &lt;!ELEMENT Entity_Node_Set ( Entity_Node+ )&gt; </pre>
--	---

```

<!ELEMENT Relation Temporal_Parallel| Temporal_Sequential |  

    Spatial_Alignment | Spatial_Arrangement |  

    Semantic_Relation | code)*>  

5   <!ATTLIST Relation  

        With_Respect_To IDREF #IMPLIED >  

  

<!ELEMENT Temporal_Parallel EMPTY >  

<!ELEMENT Temporal_Sequential EMPTY >  

10  <!ATTLIST Temporal_Sequential  

        Pattern (EXACT|DELAY|PRIOR) "EXACT" >  

  

<!ELEMENT Spatial_Alignment EMPTY >  

<!ATTLIST Spatial_Alignment  

        Pattern (Left_Align | Right_Align |  

        Top_Align | Bottom_Align) "Left_Align"  

        At_Time CDATA #IMPLIED >  

  

20  <!ELEMENT Spatial_Arrangement ( Spatial_Relevance | Spatial_Positioning )? >  

<!ATTLIST Spatial_Arrangement  

        At_Time CDATA #IMPLIED >  

<!ELEMENT Spatial_Relevance EMPTY >  

<!ATTLIST Spatial_Relevance  

        Pattern (Top_Of|Bottom_Of|Left_Of|Right_Of|  

        Upper_Left_Of|Upper_Right_Of|Lower_Left_Of|  

        Lower_Right_Of |  

        Adjacent_To | Neighboring_To | Near_By |  

        Within | Contained_In ) "Top_Of" >  

  

30  <!ELEMENT Spatial_Positioning EMPTY >  

<!ATTLIST Spatial_Positioning  

        Horizontal_Shift CDATA #IMPLIED  

        Vertical_Shift CDATA #IMPLIED >  

  

35  <!ELEMENT Semantic_Relation (Keywords* | code*)? >  

  

<!ELEMENT MM_Obj_Media_Features ( Data_Location?, Scalable_Representation?,  

    Modality_Transcoding? )>  

40  <!ELEMENT MM_Obj_Semantic_Features ( Text_Annotation?, Keywords? )>  

  

<!ELEMENT MM_Obj_Temporal_Features ( Duration? )>

```

```

<!ELEMENT Keywords ( Word*, Code* )>
<!ATTLIST Keywords
      No-Words          CDATA          #REQUIRED
      Language           CDATA          "English"
      Extraction_Manner (AUTOMATIC|MANUAL) "MANUAL">

5

<!ELEMENT Duration (
      Image_Duration?,
      Audio_Duration?,
      Video_Duration?,
      Text_Duration?,
      Synthetic_Duration? )>
<!ATTLIST Duration
      Synchronized_Overall_Duration CDATA #IMPLIED >
10
<!ELEMENT Image_Duration ( Time )>
<!ELEMENT Audio_Duration ( Time )>
<!ELEMENT Video_Duration ( Time )>
<!ELEMENT Synthetic_Duration ( Time )>
<!ELEMENT Text_Duration ( Time? | Alignment? )>
15
<!ELEMENT Alignment EMPTY>
<!ATTLIST Alignment
      with (IMAGE|AUDIO|VIDEO|SYNTHETIC) "AUDIO" >

20

<!ELEMENT Data_Location(
      Image_Location?,
      Audio_Location?,
      Video_Location?,
      Text_Location?,
      Synthetic_Location?)>
25
<!ELEMENT Image_Location (location)>
<!ELEMENT Audio_Location (location)>
<!ELEMENT Video_Location (location)>
<!ELEMENT Text_Location (location)>
<!ELEMENT Synthetic_Location (location)>
30
35

<!ELEMENT Scalable_Representation ((
      Static_Sampled?,
      Dynamic_Condensed?)>
40
<!ELEMENT Static_Sampled (
      Image_Condensed?,
      Video_Static_Pictures?,
      Audio_Clips?,
```

	Synthetic_Pictures?) >
5	<!ELEMENT Image_Condensed ( Location*, Image_Scl*, Image_Subsampling* ) >
	<!ELEMENT Image_Subsampling ( Image_Subsampling_Para, code)* >
	<!ELEMENT Image_Subsampling_Para EMPTY >
5	<!ATTLIST Image_Subsampling_Para
	Scheme       CDATA #REQUIRED
	Spatial_Rate   CDATA #REQUIRED
	Frame_Size    CDATA #IMPLIED >
10	<!ELEMENT Video_Static_Pictures (Key_Frame*) >
10	<!ATTLIST No_KFs           CDATA #REQUIRED >
	<!ELEMENT Audio_Clips (Audio_Object*   Audio_Hierarchy*) >
	<!ATTLIST No_Clips          CDATA #REQUIRED >
15	<!ELEMENT Synthetic_Pictures (Key_Frame)* >
15	<!ATTLIST No-KFs          CDATA #REQUIRED >
20	<!ELEMENT Dynamic_Condensed (
	Visual_Condensed?,
	Audio_Condensed?,
20	Text_Condensed?,
	Synthetic_Condensed?) >
25	<!ELEMENT Visual_Condensed ( Location*, Video_Scl*, Video_Subsampling* ) >
	<!ELEMENT Video_Subsampling ( Video_Subsampling_Para, code)* >
	<!ELEMENT Video_Subsampling_Para EMPTY >
25	<!ATTLIST Video_Subsampling_Para
	Scheme       CDATA #REQUIRED
	Temporal_Rate   CDATA #IMPLIED
	Spatial_Rate    CDATA #IMPLIED
	Frame_Size     CDATA #IMPLIED >
30	<!ELEMENT Audio_Condensed ( Location*,
	Audio_Compressed*,
	Audio_Subsampling*,
30	Audio_Timescaled* ) >
35	<!ELEMENT Audio_Compressed ( Audio_Compress_Para, code )* >
35	<!ELEMENT Audio_Compress_Para EMPTY >
	<!ATTLIST Audio_Compress_Para
	Scheme       CDATA #REQUIRED
	Bitrate      CDATA #IMPLIED >
40	<!ELEMENT Audio_Subsampling ( Audio_Subsampling_Para, code)* >
40	<!ELEMENT Audio_Subsampling_Para EMPTY >
	<!ATTLIST Audio_Subsampling_Para
	Scheme       CDATA #REQUIRED
	Temporal_Rate   CDATA #IMPLIED >
	<!ELEMENT Audio_Timescaled ( Audio_Timescale_Para, code)* >

```

<!ELEMENT Audio_Timescale_Para EMPTY >
<!ATTLIST Audio_Timescale_Para
      Scale_Rate      CDATA #REQUIRED >

5  <!ELEMENT Text_Condensed (Text_Abstract*) >
<!ELEMENT Text_Abstract (Location*, (Text_Abstract_Para, code)*)? >
<!ATTLIST Text_Abstract
      Length_In_Words    CDATA          #IMPLIED
      Duration_In_Seconds CDATA          #IMPLIED
10     Language        CDATA          "English"
      Generation_Mode   (AUTOMATIC|MANUAL) "MANUAL" >
<!ELEMENT Text_Abstract_Para EMPTY >
<!ATTLIST Text_Abstract_Para
      Length_In_Words    CDATA          #IMPLIED
15     Language        CDATA          "English" >

<!ELEMENT Synthetic_Condensed (Synthetic_Location*, (Synthetic_Condense_Para,
code)*) >
<!ELEMENT Synthetic_Condense_Para EMPTY >
<!ATTLIST Synthetic_Condense_Para
      Spatial_Rate       CDATA          #REQUIRED
      Temporal_Rate     CDATA          #REQUIRED
      Frame_Size        CDATA          #REQUIRED
      Bitrate           CDATA          #IMPLIED >
25
<!-- Multimedia Integration DS End -->

```